

Harmonisatie RGBZ met TMLO

Van: Arjan Kloosterboer

Datum: 16-5-2014

Ter bespreking in Expertgroep Informatiemodellen dd. 22-5-2014, agendapunt 4

In april 2013 is het 'Toepassingsprofiel Metadatering Lokale Overheden' (TMLO, versie 1) gepubliceerd door het Convent van RHC's en Nationaal Archief. In het TMLO is vastgelegd welke metagegevens bij lokale overheidsorganisaties minimaal beschikbaar moeten zijn van duurzaam te bewaren en toegankelijk te houden informatieobjecten en groepen daarvan (dossiers, series, archieven).

De betrokkenheid van de lokale overheden bij het opstellen hiervan was beperkt. Om het toepassingsprofiel gedragen te krijgen door de lokale overheden is eind 2013/begin 2014 een project uitgevoerd (in opdracht van het Nationaal Archief, Programma Arhief 2020) om het TMLO af te stemmen met vertegenwoordigers van gemeenten, waterschappen en provincies en het TMLO waar nodig aan te passen. Dit heeft onlangs geresulteerd in de publicatie van versie 1.1 van het TMLO (zie [hier](#)).

In relatie tot het RGBZ kan het TMLO beschouwd worden als een informatiemodel van 'records' (archiefbestanddelen): archieven, series, dossiers en archiefstukken. Aansluiting van beide standaarden op elkaar is dringend gewenst zodat de vereiste metadata van archiefbestanddelen probleemloos ontleend kunnen worden aan de informatie die bij de behandeling van een zaak is vastgelegd over die zaak. Zie hiervoor verder het separate document 'Standaardisatie metadatering; extract', hoofdstuk 2.

In de huidige conceptversie van het wijzigingsvoorstel RGBZ zijn al aanpassingen opgenomen ter harmonisatie met versie 1.0 van het TMLO. Harmonisatie met versie 1.1 behoeft aandacht op drie punten (zie genoemd extract, par. 3.3):

- Toevoegen van een attribuutsoort 'Gebruiksrechten' aan INFORMATIEOBJECT;
- Verschil in modellering van een samengesteld document;
- Toevoegen van bestandskenmerken aan ENKELVOUDIG INFORMATIEOBJECT.

De attribuutsoort Gebruiksrechten is reeds opgenomen in het huidige wijzigingsvoorstel RGBZ. Hieronder gaan we achtereenvolgens in op de andere twee punten. Het is de bedoeling dat we, op basis van discussie hierover, tot een voorkeur voor een oplossingsrichting komen in de komende expertgroep.

Modellering samengestelde documenten

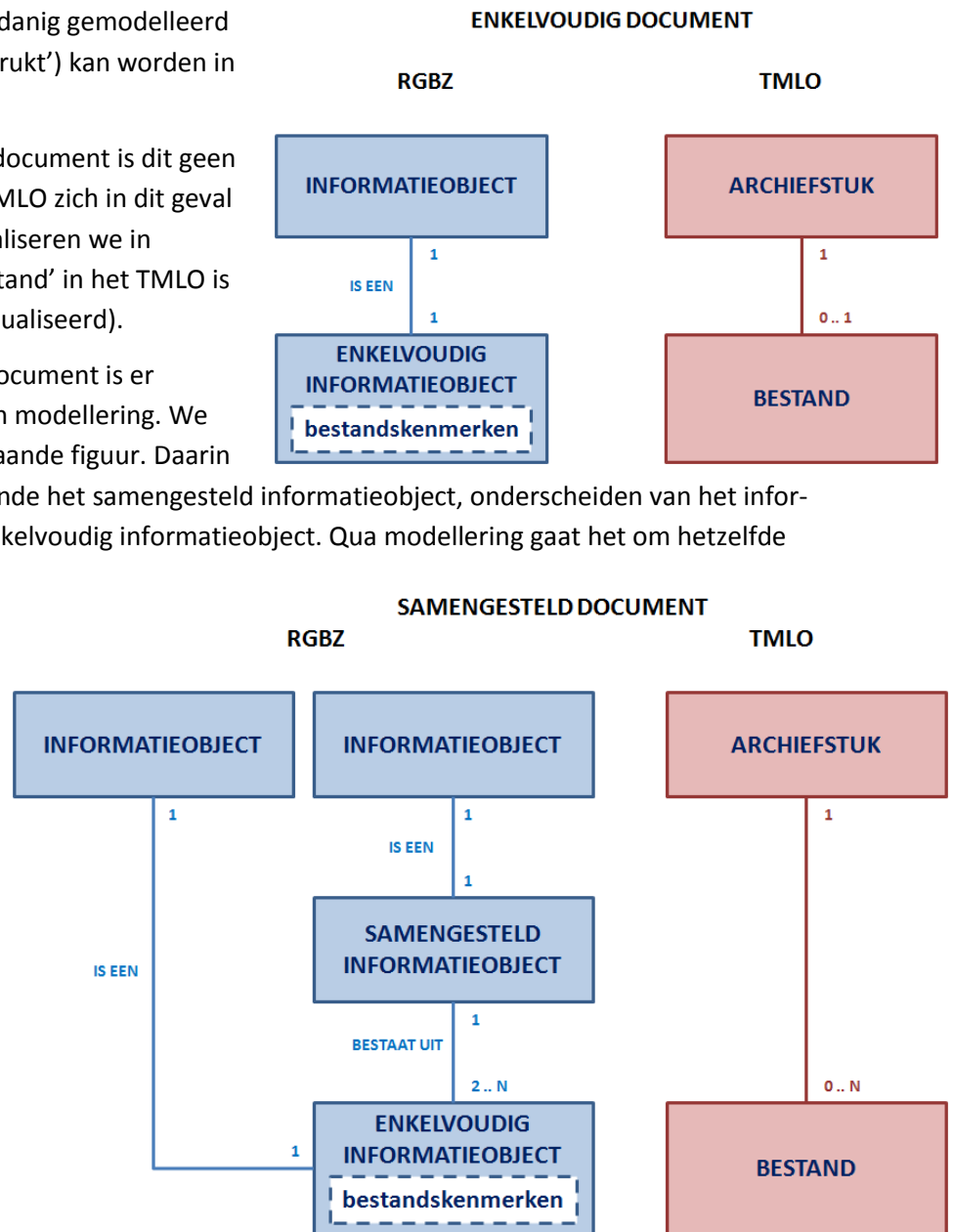
Een 'document' is in het RGBZ gemodelleerd door middel van het INFORMATIEOBJECT en de specialisaties daarvan: ENKELVOUDIG INFORMATIEOBJECT en SAMENGESTELD INFORMATIEOBJECT waarbij een SAMENGESTELD INFORMATIEOBJECT bestaat uit twee of meer ENKELVOUDIGE INFORMATIEOBJECTen. In het TMLO is een 'document' gemodelleerd als RECORD zijnde een Archiefstuk met een herhalend groepattribuutsoort Formaat waarmee telkens één bestand vastgelegd wordt dat deel uit maakt van het archiefstuk cq. document.

De vraag is of het RGBZ zodanig gemodelleerd is dat dit 'gemapt' ('uitgedrukt') kan worden in het TMLO.

V.w.b. een 'enkelvoudig' document is dit geen probleem. Hoe RGBZ en TMLO zich in dit geval tot elkaar verhouden visualiseren we in nevenstaande figuur ('Bestand' in het TMLO is daarin als objecttype gevisualiseerd).

Voor een samengesteld document is er nadrukkelijk een verschil in modellering. We visualiseren dit in onderstaande figuur. Daarin is het informatieobject, zijnde het samengesteld informatieobject, onderscheiden van het informatieobject, zijnde een enkelvoudig informatieobject. Qua modellering gaat het om hetzelfde objecttype.

Een voorbeeld van een samengesteld document is een aanvraag bouwvergunning, bestaande uit het ingevulde aanvraagformulier, een rapport met bouwtechnische berekeningen en enkele A0-bouwtekeningen. In digitale vorm gaat het dan bijvoorbeeld om twee pdf's (aanvraagformulier en rapport), een spreadsheet (met de berekening) en enkele Autocad-bestanden (met de tekeningen).



Er is een onderscheid in consequenties van de verschillen tussen een papieren en een digitaal samengesteld document.

Bij een **papieren samengesteld document** is er in het TMLO alleen sprake van een archiefstuk, niet van bestanden. Indien de documenten, waaruit het samengestelde document bestaat, als zodanig worden vastgelegd dan is er in het RGBZ sprake van een informatieobject zijnde een samengesteld informatieobject en evenzovele informatieobjecten zijnde enkelvoudige informatieobjecten als dat er documenten deel uit maken van het samengestelde document. Dergelijke (enkelvoudige) informatieobjecten kunnen alleen conform het TMLO vastgelegd worden als aparte archiefstukken. Met metadataelement 15 in het TMLO kan de relatie vastgelegd worden tussen al deze documenten. Al deze documenten hebben een eigen set aan metagegevens wat op zich overeenstemt met de modellering als aparte informatieobjecten in het RGBZ. Waarden van bepaalde metadataelementen van de enkelvoudige documenten worden overerft van het samengestelde document. In het RGBZ is

daarin niet voorzien. In het TMLO zou dit afgeleid kunnen worden vanwege de vastgelegde relaties in element 15. Het lijkt er evenwel op dat het TMLO niet gericht is op het vastleggen van de onderdelen van een papieren samengesteld document. Het RGBZ is dan goed uit te drukken in het TMLO waarbij metagegevens van de papieren enkelvoudige documenten (die deel uit maken van het samengestelde document) vastgelegd worden met element 19.3 van het TMLO (Structuur) bij het archiefstuk zijnde het samengestelde document. Deze wijze van vastlegging conform het TMLO komt dan overeen met de wijze van vastlegging van digitale samengestelde documenten. Overeenstemming op dit punt is aan te bevelen om het 'niet te ingewikkeld te maken' in de bijhouding van zaak- en documentgegevens.

Bij een **digitaal samengesteld document** is er in het TMLO sprake van een archiefstuk en van een bestand voor elk bestand dat deel uit maakt van het document. Een archiefstuk kent 'documentmetadata', een bestand alleen 'technische metadata'. In het RGBZ is er sprake van een informatieobject zijnde een samengesteld informatieobject met evenzovele enkelvoudige informatieobjecten als dat er bestanden deel uit maken van het samengestelde document. Elk enkelvoudig informatieobject is tevens een informatieobject. Het verschil met het TMLO zit 'm daarin dat ook van bestanden 'documentmetadata' worden vastgelegd (met name bij informatieobject). Bij het uitdrukken van het RGBZ in het TMLO is voor die metadata geen plaats. Enkel de 'technische metadata' kan conform het TMLO worden vastgelegd. Er treedt dan verlies van informatie op.

Er zijn globaal twee **oplossingsrichtingen** om de geconstateerde verschillen te overbruggen. De ene oplossing is om de modellering van informatieobjecten in het RGBZ overeenkomstig te maken met het TMLO. Dit betekent dat het SAMENGESTELD INFORMATIEOBJECT vervalt, dat de 'documentmetadata' van het ENKELVOUDIG INFORMATIEOBJECT worden overgeheveld naar INFORMATIEOBJECT en dat ENKELVOUDIG INFORMATIEOBJECT wordt hernoemd naar BESTAND oid. De andere oplossing is om de modellering in het RGBZ zo te houden en alleen te bekijken of er wellicht 'documentmetadata' overgehevend moet worden van ENKELVOUDIG INFORMATIEOBJECT naar INFORMATIEOBJECT.

Het nadeel van de eerste oplossing is dat geen niet-technische metadata meer vastgelegd kan worden van de documenten die deel uit maken van een samengesteld document. Dit is vooral een gemis als een samengesteld document relevant is voor de ene zaak terwijl voor een andere zaak alleen één van de onderdelen van dat samengestelde document relevant is. Het nadeel van de tweede oplossing is dat de modellering van RGBZ en TMLO v.w.b. samengestelde documenten afwijkt. Evenwel, beseft moet worden dat beide modellen hun eigen doel hebben en afwijkingen gerechtvaardigd kunnen zijn. Mits het één in het ander uitgedrukt kan worden. Van RGBZ naar TMLO is dat mogelijk, omgekeerd niet zonder meer. Gezien de levenscyclus van documenten mag verwacht worden dat van de laatste situatie geen sprake zal zijn.

Mijn voorkeur gaat er naar uit de modellering van het RGBZ te houden zoals die is, met genoemde kanttekeningen.

Toevoegen van bestandskenmerken aan ENKELVOUDIG INFORMATIEOBJECT

Het TMLO kent van elk bestand, dat deel uit maakt van een archiefstuk, meerdere metagegevens.

Een dergelijk bestand komt overeen met het ENKELVOUDIG INFORMATIEOBJECT in het RGBZ.

Idealiter zijn in het TMLO opgenomen bestandskenmerken ook gemodelleerd in het RGBZ. Hieronder sommen we de bestandskenmerken in het TMLO op met daarachter het overeenkomende attribuut in het RGBZ bij het ENKELVOUDIG INFORMATIEOBJECT.

TMLO-element	RGBZ-attribuut	Opmerking
21.1 Identificatiekenmerk	Bestandsnaam ?	Betreft in het TMLO de unieke sleutel waaronder het bestand eenduidig gevonden kan worden op de fysieke of virtuele locatie zoals aangeduid met element 7: Plaats. Dit kan de bestandsnaam zijn mits deze uniek is of, door middel van voorloopkenmerken of aanvullende kenmerken (zoals een 'pad-beschrijving), uniek gemaakt is.
21.2 Naam		
○ 21.2.1 Naam	Titel (of Bestandsnaam?)	Betreft in het TMLO de naam van het bestand, zonder de formaat-extensie (zoals .pdf), waarmee de inhoud van het bestand gekenschetst wordt.
○ 21.2.2 Extensie	Formaat	Betreft in het TMLO de aanduiding van het bestandsformaat
21.3 Type	"enkelvoudig"	Of ook rekening houden met samengestelde, enveloppe en container bestanden?
21.4 Omvang	Bestandsomvang	
21.5 Bestandsformaat		Betreft in het TMLO de code volgens welke gegevens op een gegevensdrager zijn opgeslagen conform de bestandsformaat-aanduidingen zoals benoemd in het PRONOM-register, bijvoorbeeld "fmt/40" (Microsoft Word 97-2003 document; extensie: .doc). De bestands-extensie van element 21.2.2. hoeft niet overeen te komen met het eigenlijke bestandsformaat en verschaft onvoldoende informatie.

21.6 Creatieapplicatie		Omschrijving van de applicatie waarmee het bestand gemaakt is.
○ 21.6.1 Naam		
○ 21.6.2 Versie		
○ 21.6.4 Datum aanmaak	Begindatum Materiele historie van Formaat	
21.7 Fysieke integriteit		Uitdrukking van mate van volledigheid en onbeschadigd zijn van digitaal bestand.
○ 21.7.1 Algoritme (type)		
○ 21.7.2 Waarde		
○ 21.7.3 Datum		
21.8 Datum aanmaak	Begindatum Materiele historie van Bestandsnaam	
21.9 Event plan formaat	-	Speelt pas een rol bij archiefbeheer, niet bij zaakbehandeling
○ 21.9.1 Datum	-	
○ 21.9.2 Type	-	
○ 21.9.3 Beschrijving	-	
○ 21.9.4 Aanleiding	-	
21.10 Relatie		N.v.t. bij Type (21.3) "enkelvoudig". Zie opm aldaar.

Onduidelijkheid is er over 'Identificatiekenmerk' (21.1) en 'Naam' (21.2.1.). Passen we het RGBZ op deze punten aan op het TMLO? Mijn voorstel is: ja.

V.w.b. de elementen 'Type' (21.3) en 'Relatie' (21.10) is de vraag of we met het RGBZ alleen enkelvoudige bestanden ondersteunen (een bestand dat niet bestaat uit andere bestanden) of ook bestanden die opgebouwd zijn uit andere bestanden (bijv. een zip-file). Mijn voorstel is: ja.

Verder zijn er enkele elementen in het TMLO die niet voorkomen in het RGBZ en waarvan we moeten overwegen deze ook op te nemen in het RGBZ. Dat betreffen Bestandsformaat (21.5), Creatieapplicatie (21.6) en Fysieke integriteit (21.7). Ik stel voor deze elementen aan het RGBZ toe te voegen (optioneel te vullen)

Al met al behelst het voorstel maximale aansluiting van het TMLO op het RGBZ teneinde documenten consistent te kunnen metadateren in de gehele levenscyclus met een zo gering mogelijke beheerinspanning.